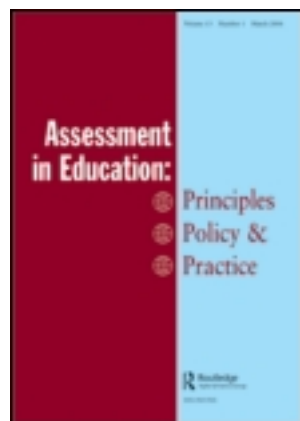


This article was downloaded by: [University of Manitoba Libraries]

On: 16 April 2013, At: 16:43

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Assessment in Education: Principles, Policy & Practice

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/caie20>

### Teachers' conceptions of assessment: implications for policy and professional development

Gavin T. L. Brown<sup>a</sup>

<sup>a</sup> University of Auckland, New Zealand

Version of record first published: 27 Sep 2010.

To cite this article: Gavin T. L. Brown (2004): Teachers' conceptions of assessment: implications for policy and professional development, *Assessment in Education: Principles, Policy & Practice*, 11:3, 301-318

To link to this article: <http://dx.doi.org/10.1080/0969594042000304609>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# Teachers' conceptions of assessment: implications for policy and professional development

Gavin T. L. Brown<sup>\*</sup>

*University of Auckland, New Zealand*

Teachers' conceptions of assessment can be understood in terms of their agreement or disagreement with four purposes to which assessment may be put, specifically, (a) improvement of teaching and learning, (b) school accountability, (c) student accountability, or (d) treating assessment as irrelevant. A 50-item Teachers' Conceptions of Assessment (COA-III) questionnaire was completed by New Zealand primary school teachers and managers ( $n=525$ ). The COA-III, based on the four main purpose-defined conceptions of assessment, was analysed with structural equation modelling and showed a close fit of the data to a hierarchical, multi-dimensional model ( $\chi^2=3217.68$ ;  $df=1162$ ;  $RMSEA=.058$ ;  $TLI=.967$ ). On average, participants agreed with the improvement conceptions and the school accountability conception, while rejecting the view that assessment was irrelevant. However, respondents disagreed that assessment was for student accountability. Improvement, school, and student accountability conceptions were positively correlated. The irrelevance conception was inversely related to the improvement conception and not related to the system accountability conception. Surprisingly, no statistically significant differences were found in mean scale scores for each conception regardless of teacher (age, gender, role, assessment training, or assessment practices) or school (size, location, or socio-economic status) variables. Implications for the use of the COA-III for policy implementation and teacher professional development are discussed.

National, district, or school policies that involve the professional identity of teachers do not usually correspond to the opinions and conceptions of teachers (van den Berg, 2002). Furthermore, many policies concerning assessment standards and procedures aim to connect teaching and learning to regulation and administration. Thus, the success or failure of such policies may hang on the conceptions and meanings that teachers give to those policies. Additionally, the impact of professional development may be enhanced if greater attention were given to the identification of teachers' meanings or understandings (Hawley & Valli, 1999; Cohen & Hill, 2000).

---

<sup>\*</sup>School of Education, University of Auckland, Private Bag 92019, Auckland, New Zealand.  
Email: [gt.brown@auckland.ac.nz](mailto:gt.brown@auckland.ac.nz)

For example, the following hypothetical scenario involves five teachers (A–E) reacting to a notice advertising a new assessment package intended to promote a national policy initiative of improving teachers' assessment literacy.

**A:** See! All they're interested in is checking up on us. How can they keep using tests to decide if we're good teachers or not? What's the union doing to protect us?

**B:** Why worry? Tests are there to find out if students are good at school work—you know just intelligence tests. Our kids will only do well if they study and practice what we teach them; if they don't then it's their own fault they don't pass, not ours. Nobody can blame us for our kids' results.

**C:** That might be, but you know what to do, don't you? If they make you use it, just do it, write the scores down and forget about it and carry on doing what you always do. After all we're good teachers; we know what our kids are like and what they need. We don't need any tests to help us do a good job!

**D:** I'm not so sure about that. I've seen the trial stuff when our kids did it last year. The kids in my class really enjoyed them—it made them work a little harder and feel good about themselves. I think this kind of assessment might just motivate our kids.

**E:** Well, I've seen them too and I think the reports will help us do our jobs better. There are all kinds of descriptive information in them about what achievement objectives kids need to work on, what their strengths are, and what they've already mastered. It gives you all sorts of good ideas about where to start and who needs what.

In this hypothetical conversation, four major conceptions of assessment can be derived: (1) assessment is related to improvement of student learning and teachers' instruction (teachers D and E); (2) assessment makes students accountable for learning (teacher B); (3) assessment evaluates the quality of schools and teachers (teacher A); and (4) assessment is irrelevant to the work of teachers (teacher C). It is the goal of this paper to show that the existence of these four conceptions in the minds of New Zealand primary school teachers is not just a hypothetical assertion but a reality.

### **Teachers' conceptions**

Teachers' beliefs are organized into systems wherein some beliefs are more central or primary while others are peripherally linked to those central beliefs (Pajares, 1992; Thompson, 1992). Beliefs are the meanings connected to psychological objects or phenomena and are an environmentally contingent and culturally defined lens through which sense is made of events, people, and interactions (Pratt, 1992; Ekeblad & Bond, 1994). A wide variety of language has been used to refer to teachers' beliefs, including 'teachers' subjectively reasonable beliefs' (Harootunian & Yarger, 1981), 'untested assumptions' (Calderhead, 1996), and 'implicit theories' (Clark & Peterson, 1986).

However, to address the varying terminology about knowledge, beliefs, belief systems, and belief clusters more efficiently, Thompson (1992) invoked concep-

tions 'as a more general mental structure, encompassing beliefs, meanings, concepts, propositions, rules, mental images, preferences, and the like' (p. 130). Furthermore, conceptions represent different categories of ideas held by teachers behind their descriptions of how educational things are experienced (Pratt, 1992). Thus, conceptions act as a framework through which a teacher views, interprets and interacts with the teaching environment (Marton, 1981). 'Conceptions' is the term used in this research to describe the organizing framework by which an individual understands, responds to, and interacts with a phenomenon. The structure of teachers' conceptions is not uniform and simple; they appear to be multifaceted and interconnected.

The study of teachers' conceptions of assessment is important because evidence exists that teachers' conceptions of teaching, learning, and curricula influence strongly how they teach and what students learn or achieve (Clark & Peterson, 1986; Pajares, 1992; Thompson, 1992; Calderhead, 1996). Indeed, teachers' beliefs about student self-confidence, morale, creativity, and work are 'closely linked to one's choice of evaluation techniques' (Asch, 1976, p. 18). Tittle (1994) proposed that teachers 'construct schemas or integrate representations from assessments into existing views of the self, of teaching and learning, and of the curriculum, broadly construed' (p. 151). From their survey of elementary school teachers, Cizek *et al.*, (1995) argued that, based on the highly individualistic nature of assessment practices, many teachers seem to have assessment policies based on their idiosyncratic values and conceptions of teaching. In a study of high school English classes, Kahn (2000) has argued that teachers used a wide variety of seemingly conflicting assessment types because they eclectically held and practised transmission-oriented and constructivist models of teaching and learning. And yet, as individualistic as conceptions may appear, it can be argued they are socially and culturally shared cognitive configurations or phenomena (van den Berg, 2002).

Thus, all pedagogical acts, including teachers' perceptions and evaluations of student behaviour and performance (i.e., assessment), are affected by the conceptions teachers have about many educational artefacts, such as teaching, learning, assessment, curriculum, and teacher efficacy. It is critical that such conceptions and the relationships of those conceptions among and between each other, are made explicit and visible. This is especially so if it is considered prudent or advisable that teachers' conceptions be changed, which, of course, is the point of professional development activities (Borko *et al.*, 1997).

This article describes four conceptions of assessment that teachers may hold, and reports the empirical results of a survey aimed at identifying the socially and culturally shared conceptions of New Zealand primary school teachers about assessment. Further, on the assumption that conceptions act as filters through which teachers interpret and experience state sponsored or school-wide policies and practices and even professional development activities related to assessment, the article concludes with some possible implications for assessment policy and assessment-related professional development.

#### Four conceptions of assessment

Assessment is any act of interpreting information about student performance, collected through any of a multitude of means or practices. Assessment, according to the Department of Education in England (as cited in Gipps *et al.*, 1995) involves

a broad appraisal including many sources of evidence and many aspects of a pupil's knowledge, understanding, skills and attitudes; or to a particular occasion or instrument....any method or procedure, formal or informal, for producing information about pupils: e.g., [sic] a written test paper, an interview schedule, a measurement task using equipment, a class quiz. (pp. 10–11)

Researchers have suggested that three major purposes for assessment exist: improvement of teaching and learning, making students accountable for learning partly through issuing certificates, and accountability of schools and teachers (Heaton, 1975; Webb, 1992; Torrance & Pryor, 1998; Warren & Nisbet, 1999). In addition to the three conceptions, a fourth conception can be expected; that is, assessment is fundamentally irrelevant to the life and work of teachers and students perhaps because it is bad for teachers and students, or because it can be safely ignored even if it must be used, or even because it is inaccurate.

The major premise of the improvement conception is that assessment improves students' own learning and the quality of teaching (Crooks, 1988; Black & Wiliam, 1998). This improvement has two important caveats; (a) assessment must describe or diagnose the nature of student performance and (b) the information must be a valid, reliable, and accurate description of student performance. In this view, a range of techniques, including informal teacher-based intuitive judgement as well as formal assessment tools, identify the content and processes of student learning, including impediments to learning and unexpected strengths, with the explicit goal of improving the quality of instruction and student learning.

A second conception of assessment is that assessment can be used to account for a teacher's, a school's, or a system's use of society's resources (Firestone *et al.*, 1998). This conception uses assessment results to demonstrate publicly that teachers or schools are doing a good job (Mehrens & Lehmann, 1984; Butterfield *et al.*, 1999; Smith *et al.*, 1999) and imposes consequences for schools or teachers for reaching or not reaching required standards (Firestone *et al.*, 1998; Guthrie, 2002). Two rationales for this conception exist; one emphasizes demonstrating publicly that schools and teachers deliver quality instruction (Smith & Fey, 2000; Hershberg, 2002), and the second emphasizes improving the quality of instruction (Noble & Smith, 1994; Linn, 2000).

The premise of the third conception of assessment is that students are held individually accountable for their learning through assessment. This is seen in the assignment of grades or scores, checking off student performance against criteria, placing students into classes or groups based on performance, as well as various qualifications examinations in which secondary age students participate for graduation or entry selection to higher levels of educational opportunity. In New Zealand primary schools, the use of assessment for student accountability focuses much more

on determining whether students have met various curriculum objectives (Hill, 2000), the criteria for a given curriculum level (Dixon, 1999), or merit placement in a certain learning group within a class. The certification of students in New Zealand is largely a secondary school activity during the final three years of schooling and there are many significant consequences for individuals dependent on their performance on such assessments, including retention in a year or grade level, graduation, and tracking or streaming (Guthrie, 2002). Together, these uses instantiate a conception wherein assessment is used as a means of making students accountable for learning.

The premise of the final conception is that assessment, usually understood as a formal, organized process of evaluating student performance, has no legitimate place within teaching and learning. Teachers' knowledge of students based on long relationship and their understanding of curriculum and pedagogy preclude the need to carry out any kind of assessment beyond the intuitive in-the-head process that occurs automatically as teachers interact with students (i.e., Airasian's (1997) 'sizing up'). Assessment may be rejected also because of its pernicious effects on teacher autonomy and professionalism and its distractive power from the real purpose of teaching (i.e., student learning) (Dixon, 1999). Teachers of English in England welcomed a new National Curriculum in the early 1990s but rejected the accountability assessments because the Key Stage assessments were considered inimical to the learning and teaching values espoused in the curriculum (Cooper & Davies, 1993). It may also be that the degree of inaccuracy (e.g., standard error of measurement) published with any formal assessment contributes to teachers' conception of assessment as irrelevant.

These purposes can lead to different practices, and often there can be tensions between the purposes. For example, the tension between externally imposed accountability requirements and the improvement conception has created difficulties for New Zealand teachers (Dixon, 1999; Hill, 2000). The various conceptions might also interact with each other. For example, teachers who see assessment as irrelevant could also possibly believe that improvement is the legitimate goal of teacher judgement and may yet reject assessment as a legitimate means of reaching that goal. On the other hand, concern for improvement may associate strongly with school self-managed accountability but less strongly with the student accountability view. Awareness of measurement error in assessment may lead to an irrelevance view of assessment. It was the goal of this research to determine whether these purposes exist in the minds of teachers and how they may be structured.

Thus, this article asserts that teachers' conceptions of assessment constitute four major beliefs about assessment: (a) assessment improves teacher instruction and student learning by providing quality information for decision-making; (b) assessment makes students accountable for their learning; (c) teachers or schools are made accountable through assessment; and (d) assessment is irrelevant to the work of teachers and the life of students.

The empirical research on teachers' conceptions of assessment, as opposed to their observed or reported assessment practices (e.g., Gipps, *et al.*, 1995; Quilter, 1998; Dixon, 1999; Hill, 2000; McMillan *et al.*, 2002) or the literature advising teachers



how to use assessment (e.g., Mehrens & Lehmann, 1984; Airasian, 1997; Linn & Gronlund, 2000; Popham, 2000; McMillan, 2001), is limited and such a review is being prepared for publication elsewhere. Before reporting this study of New Zealand primary school teachers' conceptions of assessment, a brief outline of the New Zealand primary school assessment practices and policies is given.

### **New Zealand context**

In the last two decades, as in other jurisdictions, large structural changes have been initiated in New Zealand schooling and education (Fiske & Ladd, 2000; Levin, 2001; Crooks, 2002). The present New Zealand Ministry of Education (MoE) is a policy only body; while other statutory bodies deal with important functions devolved from the MoE; specifically, the Education Review Office (ERO) was made responsible for quality assurance of schools, the New Zealand Qualifications Authority (NZQA) was made responsible for secondary and tertiary level qualifications. The New Zealand Curriculum Framework (NCF) consists of seven essential learning areas (i.e., Language and Languages, Mathematics, Science, Social Science, Physical Well-being and Health, Technology, and Arts) each of which has eight hierarchical levels of achievement covering Years 1–13 (primary and secondary schooling) (Ministry of Education, 1993). The goal of NCF policy developments was a seamless education system that wove together curriculum and qualifications from childhood to adulthood.

Perhaps the most radical of governance reforms was the making of all schools responsible for their own administration and management, through single-school boards (Wylie, 1997). This means that each of the approximately 2200 primary schools in New Zealand is by legislation self-governing and self-managing, including responsibility for the selection, employment, and further professional development of its own staff and for setting policies within the school to meet Ministry mandated administrative guidelines and educational goals. To balance this relatively free hand, the government has mandated accountability inspections conducted by the Educational Review Office to verify that schools were complying with this legislation. In addition, legislation (the National Educational Goals and National Administrative Guidelines) was enacted that required schools to ensure that students reached expected levels of achievement, especially in literacy and numeracy.

The Ministry's national policy in the primary school sector emphasizes voluntary, school-based assessment for the purpose of raising achievement and improving the quality of teaching programmes (Ministry of Education, 1994). There is no compulsory state mandated assessment regime and so all assessment practices are voluntary and low stakes. In the context of self-managing schools, assessment practices are school-based. At the time of the reforms, high proportions of schools reported use in at least one class of the voluntary, standardized *Progressive Achievement Tests* of language skills while half reported using the same series' mathematics tests (Croft & Reid, 1991). More recently, it was found that a large number of standardized achievement and diagnostic assessment tools were being used in New Zealand

primary schools, with most teachers reporting that the use of voluntary diagnostic assessments frequently or always altered the way they taught their students (Croft *et al.*, 2000). The stake or consequence of these school-based assessments is low with teachers assessing their students to identify their strengths and weaknesses in progress towards curriculum objectives and to evaluate the quality of teaching programmes (Hill, 2000).

A concomitant policy on assessment within primary schools is that it ought to provide clear indicators to all concerned of student performance relative to the outcomes specified in the national eight-level curriculum statements. National testing of primary-age children against the New Zealand standards-based, eight-level curriculum has been mooted, especially at key transition points within the system (Ministry of Education, 1994; New Zealand, 1998). However, unlike England or Australia, such national assessment schemes have not been adopted in New Zealand (Levin, 2001); rather voluntary-use nationally standardized assessment tools (e.g., exemplars, item resource banks, computerised teacher-managed testing tools) have been provided to teachers (Crooks, 2002). Notwithstanding the devolution of professional development responsibility, the Ministry has also funded specialized programmes that focus on improving teachers' use of assessment for improved learning (e.g., Assessment for Better Learning, Assess to Learn). Crooks (2002) provides further details of the secondary sector assessment policy and system for interested readers.

Thus, research into New Zealand primary school teachers' conceptions of assessment takes place in a policy and practice context of self-managed, low-stakes assessment for the purpose of improving the quality of teaching and learning. Simultaneously, schools are expected to report student performance against the objectives of various curriculum statements to parent communities, while central agencies seek to obtain evidence and surety that students and schools are meeting expected standards and outcomes. This objective has been assisted by the introduction by the Ministry of various national assessment tools and training innovations focused on assessment for learning and by a continued resistance to traditional forms of national testing. This stands in some contrast to the secondary school context where centrally administered, high-stakes qualifications assessment takes place largely to determine whether students meet various standards.

### **New Zealand primary school teachers' conceptions of assessment**

A self-report attitude inventory (COA-III) about teachers' conceptions of assessment consisting of 65 statements was used. Teachers identified the degree to which they agreed or disagreed with each statement using a 6-point positively-packed agreement-rating scale consisting of two negative and four positive rating points (see Brown, 2004, for a discussion of the merits of this rating scale). Each statement expressed an opinion about assessment derived from the literature and had been subjected to exploratory factor analysis based on responses of education students and practising teachers. Those two earlier versions of the questionnaire (i.e., COA-I and COA-II)



Table 1. Expected factors of New Zealand primary school teachers' conceptions of assessment

Second Order Factor	First Order Factor
Improvement	Assessment provides valid information
	Assessment describes student learning
	Assessment improves the quality of teaching
	Assessment improves student learning
	Assessment is inaccurate
Irrelevance	Assessment is bad for students and teachers
	Assessment is used but ignored
—	Assessment makes schools accountable
—	Assessment makes students accountable

and two different participant populations identified poor fitting statements and gave early indication about the beliefs of New Zealand primary school teacher trainees and teachers (see Brown, 2002, for details). Based on the literature review and two rounds of data analysis, a multi-factoral, two-level, inter-correlated model of teachers' conceptions of assessment was developed in which the statements created nine conceptions of assessment, seven of which load onto two higher-order factors (see Table 1).

The technique used to determine whether NZ primary school teachers' conceptions of assessment are represented by the proposed model, how strongly those conceptions might be held, and how those conceptions might interrelate is structural equation modelling (SEM). SEM is a sophisticated correlational technique, sometimes known as confirmatory factor analysis, utilizing large data sets (usually >500 participants) to detect and explain relationships among meaningfully related structures (Maruyama, 1998). In other words, the specification of an SEM model is dependent on theory and ought to be used only in the context of theoretically determined relationships. Because of this power, it is able to go beyond describing individual teachers' conceptions of assessment to establishing the relationships (both strength and direction) between the various conceptions of assessment of social groups. Measurement models evaluated with SEM may contain first and second order factors representing the meaningful structures derived from the questionnaire statements to which teachers indicated their degree of agreement.

Unlike exploratory factor analysis which leads to theory based on observed data, SEM is a more conservative and powerful approach that determines the degree to which data fit the theoretically expected relationships. Additionally, SEM measures the degree and direction of correlation between factors by taking into account all correlations and covariances among all items in the data matrix simultaneously. Traditional factor analysis techniques depend on within-factor inter-correlations (e.g., alpha correlation estimates of reliability) to give confidence to the interpretation that the observed data are measuring something in common. In SEM analysis, the

overall fit of all the factors in the model is best expressed by the Root Mean Square Error of Approximation (RMSEA) which should be below .08 (Hoyle, 1995).

Two COA-III questionnaires were sent in the last half of the 2001 school year to the principals of 800 randomly, but representatively, selected primary schools with the request that one teacher and one leader/administrator of Year 5–7 students in the school be asked to complete the anonymous questionnaire voluntarily. Note that since half of all New Zealand primary schools have less than 150 students, there may have been only one or two such teachers in any school. Without any follow-up requests or inducements, teachers from 290 schools provided 491 COA-III questionnaires, while a further 34 questionnaires were returned without school identification. This sample represents between 36 and 40% of schools invited (depending on whether each of the non-identified participants represented one school) and about 33% of the maximum possible invited teachers. The participants in the study were (a) New Zealand European (83%), (b) female (76%), (c) highly experienced with 10 or more years teaching (63%), (d) about equally distributed as teachers and managers or senior teachers (54% teachers), (e) employed in contributing or full primary schools (89%), and (f) well trained with three or more years training (55%). Three key demographic characteristics of the participants in this study reasonably reflect those of the New Zealand teaching population who were 87% NZ European ethnicity, 71% female, and 50% long service in 1998 (Sturrock, 1999). Thus, data in this study were from a relatively homogenous sample of New Zealand primary school teachers and sufficiently representative of the New Zealand population of primary school teachers on which to base generalizations.

## Results

Of the 65 statements in the COA-III questionnaire, 50 items resulted in a well fitting measurement model ( $\chi^2=3217.68$ ;  $df=1162$ ;  $RMSEA=.058$ ;  $TLI=.967$ ) containing, as predicted, four correlated major factors, which constitute the conceptions of irrelevance, improvement, school accountability, and student accountability (see Figure 1). The first two factors are second-order factors that have three or four first-order factors, while the latter two are stand-alone first-order factors. The figure displays the loadings of each statement on its respective factor, the loadings of the first-order factors on the second-order factors, and the inter-correlations between the four main conceptions.

The School Accountability factor is made up of six statements that focus on using assessment to evaluate the worth of schools; while the Student Accountability conception is formed by seven statements that focus on using assessment to evaluate student progress against achievement objectives and to make placement or selection decisions about students. The first-order factor loadings for Improvement consist of four conceptions, each loading strongly on the second-order factor. These are six statements that show assessment describes student abilities, knowledge, and thinking, seven statements that indicate assessment improves student learning, six statements that point to assessment improving teaching, and five statements that demonstrate

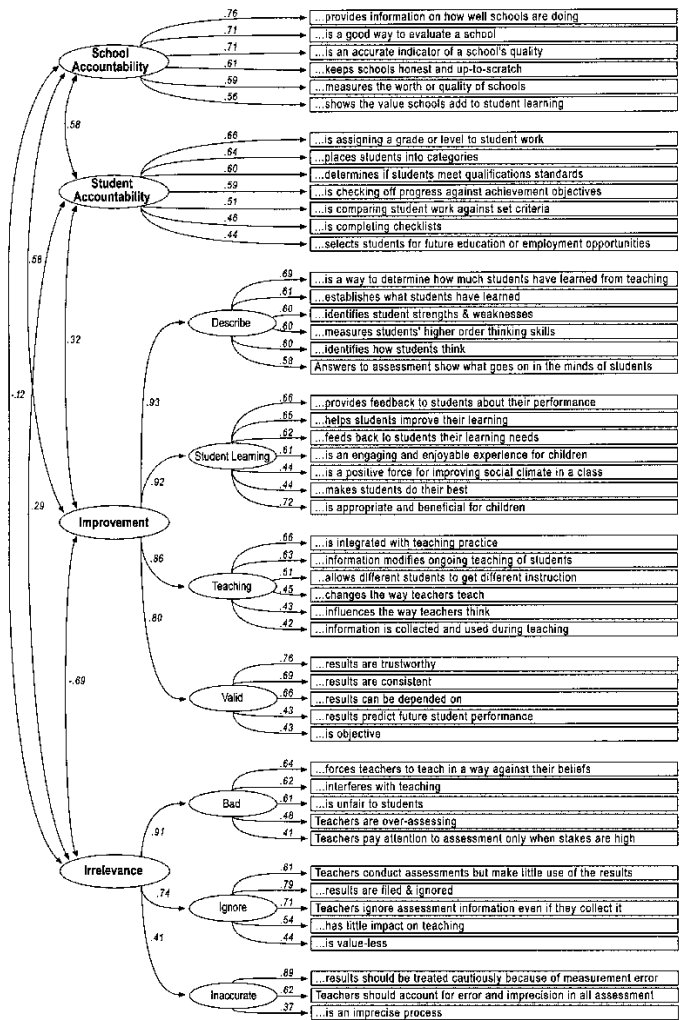


Figure 1. COA-III measurement model of conceptions of assessment

assessment information is valid because of its dependability. The first-order factor loadings for Irrelevance clearly indicate that the irrelevance conception consists of three strongly loading conceptions; specifically, five statements indicating assessment is bad for teaching, another five statements showing that teachers may use assessment but they ignore it, and three statements highlighting that assessment is inaccurate.

Note that an alternative model wherein the inaccurate factor loaded on the improvement conception was tested on the basis that cautious interpretation of assessment information is an essential aspect of using assessment to improve teaching and learning. This alternative model had significantly poorer fit than the reported model confirming that, at least in the minds of these NZ primary teachers, inaccuracy of assessment is conceived as part of irrelevance and may constitute a partial justification for ignoring assessment or considering it to be bad for students.

It is possible that any results are not generalizable or stable across various teacher or school characteristics. To eliminate these confounds, a series of multiple analysis of variance (MANOVA) studies were conducted for statistically significant differences in mean scores for the nine factor scores across the teacher characteristics of: (a) role (teacher  $n=281$  versus manager or leader  $n=218$ ); (b) years of experience (ten years or less  $n=180$  versus more than ten years  $n=319$ ), and (c) years of teacher training (less than two years  $n=107$ , two to three years  $n=116$ , three years  $n=136$ , and more than three years  $n=140$ ). The only statistically significant difference, based on the linearly independent pair-wise comparisons among the estimated marginal means, was for the improve student learning subscale,  $F(1,501)=11.691$ ,  $p=.001$ , with school leaders ( $M=4.01$ ) agreeing more strongly than teachers ( $M = 3.74$ ) that improving student learning defined assessment. This particular distinction in attitude between leaders and teachers, though it may be an artefact of experiment-wise statistical testing, has been found in the literature on teachers' implicit theories about teaching (Clark & Peterson, 1986). That only one of the nine COA-III subscales had statistically significant difference suggested that differences in role are not powerful in shaping teachers' conceptions of assessment. Thus, teacher gender, years of training, years of experience, and role in school were irrelevant to mean scale scores on the teachers' conceptions of assessment inventory.

The characteristics of the schools in which the participants worked were also eliminated as sources of statistically significant differences in mean COA-III scale scores. To permit analysis of reasonably comparable cell sizes, school socio-economic status (SES) was collapsed into three categories (i.e., low, medium, and high) and school community population size was collapsed into two categories (i.e., urban and rural). MANOVA studies of mean COA-III scale scores found no statistically significant differences (i.e., School SES,  $F(18, 434)=1.207$ , Wilks'  $\lambda=.947$ ,  $p=.248$ ; School Size,  $F(18, 434)=1.047$ , Wilks'  $\lambda=.954$ ,  $p=.403$ ; Community Type,  $F(18, 434)=1.064$ , Wilks'  $\lambda=.976$ ,  $p=.389$ ) for main or interaction effects. Thus, school characteristics did not differentiate in a statistically significant way the mean scores for the nine teachers' conceptions of assessment factors.

Another possible confounding explanation is what the word assessment actually meant to teachers. As a means of triangulating the actual wording of the statements in the COA-III, which elicited teachers' beliefs about the meaning of assessment in and of themselves, teachers were asked to identify from a list of 11 different assessment methods which ones they had in mind as they thought of the word assessment. Multi-dimensional scaling (MDS), using Alscal Euclidean distances procedure, reduced the 11 types of assessment to four meaningful dimensions, with good fit characteristics (Kruskal's stress=.026;  $R^2=.997$ ). The first dimension consisted of five types of assessment grouped together as teacher controlled classroom assessments (i.e., teacher-made written tests, standardized tests, student written work, planned observations, and student self or peer assessments) with 469 participants selecting an average of 4.03 of these types. The second was based on two formal examination types (i.e., 1–3 hour exams and essay tests) with only 61 participants selecting one or more of these types. The third dimension consisted of three oral assessments (i.e.,

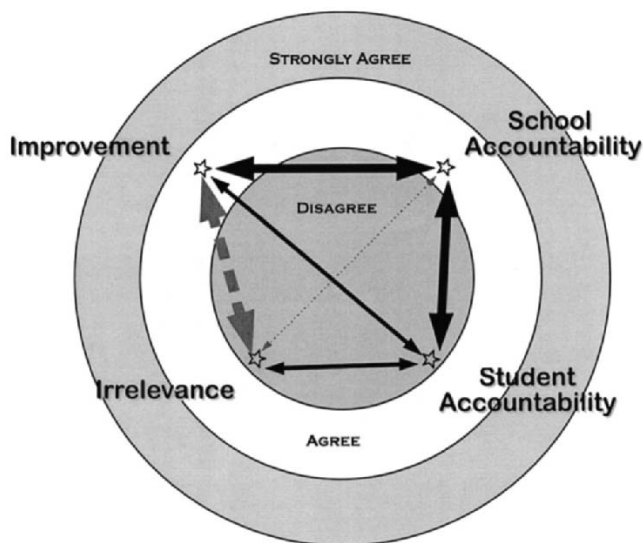


Figure 2. Strength and inter-correlations of COA-III conceptions of assessment

conferencing, oral question and answer, and unplanned observations) with 450 participants selecting an average of 1.55 of these types. The fourth dimension was the portfolio method and was selected by 311 of the participants. Thus, when thinking of assessment it was clear that these participants associated the term with teacher controlled classroom tasks, oral, and portfolio assessments—a result highly consistent with the most recent survey of NZ primary teacher assessment practices (Dunn *et al.*, 2003). This result is especially powerful in that it shows that participants were thinking of a wide variety of assessment tasks and clearly did not associate the term with formal, external examinations, as may be the case in other jurisdictions or other sectors of education. MANOVA tests revealed no statistically significant differences in mean scores for the nine conceptions of assessment scales for main or two-way interaction effects according to the four different assessment types. In other words, teachers' conceptions of assessment were general and constant and clearly related to what are commonly seen as classroom assessment tasks.

These analyses indicated the relative stability and generalizability of teachers' assessment conceptions regardless of school or teacher demographic characteristics. Not only was this study conducted with a relatively homogeneous sample, but the views of those teachers were remarkably similar. The analyses also show that the results of this study cannot be explained away by the assertion that teachers understood the word 'assessment' in some negative or inappropriate fashion. Having established that the model had good fit to the data and that it was generalizable and robust, the structural relationships of the four main conceptions in the model and the differing levels of agreement or support that teachers have for each conception were examined.

The strength of agreement for each of the four main conceptions and the intercorrelation between them provides the greatest insight into teachers' conceptions of assessment (see Figure 2). The figure shows the mean agreement score for each of the four higher-order conceptions. The concentric rings of the bull's eye show the levels of agreement with greater distance from the centre of bull's eye indicating stronger agreement. Teachers agreed with the Improvement and School Accountability conceptions and disagreed with the Irrelevance and the Student Accountability conceptions. Note that the average agreement for any conception did not reach strong agreement.

The thickness of the arrows shows the degree of inter-correlation of conceptions, while the solid lines indicate positive correlations and dashed lines showing negative correlations. Note that all correlations are statistically significant except that between Irrelevance and School Accountability conceptions. Note also the strong inverse relationship between Irrelevance and Improvement that needs to be understood as a strong positive correlation between Improvement and *Relevance*.

If teachers think assessment is about Improvement then it is unlikely they will consider assessment as Irrelevant ( $r = -.69$ ) and they are likely to believe that assessment is connected to School Accountability ( $r = .58$ ). This unexpected relationship may be because of the impact of self-management of New Zealand schools wherein teachers are accountable to their colleagues and to a school-based Board of Trustees made up of parents of pupils for the effectiveness of their work in changing student learning outcomes. Teachers who conceive of assessment as Improvement tended to have just moderate likelihood of agreeing that assessment is about Student Accountability (i.e., certifying student performance or achievement). This may be because of the impact of student-centred philosophies or conceptions.

If teachers think assessment is about School Accountability, then they may or may not believe that assessment is Irrelevant; belief in one is independent of belief in the other. Teachers who believe in assessment as School Accountability are highly likely to also conceive of assessment as a process of Student Accountability and Improvement. This suggests a nexus of conceptions around the idea that assessment for school accountability may lead to a raising of educational standards that will in turn lead to improved ability of students to receive qualifications and recognition of achievement. This is what some advocates of high-stakes accountability testing have argued would and should happen (e.g., Resnick & Resnick, 1989). However, it is worth noting that this effect is found in the New Zealand context where there are no externally mandated national tests, just a programme of school-based policies on assessment for school-based management and information.

When teachers think assessment is about Student Accountability, it is moderately likely they will also consider assessment to be Irrelevant, because it is bad for students or inaccurate, such that they can safely ignore it. It is possible that this conception is related to strong student-centred learning beliefs or humanistic curriculum or nurturing teaching beliefs. Teachers who conceive of assessment as Student Accountability are likely to have only a weak relationship to Improvement. In other words, assessment of students is likely to be Irrelevant when it is connected to Student



Accountability but is more likely to be acceptable if it is related to Improvement of teaching and learning.

When assessment is considered Irrelevant, it is highly likely to be disconnected from the goal or improving instruction or learning. This discontinuity may be driven by a rejection of Student Accountability uses of assessment, whereas it does not appear to be related at all to the conception of using assessment to evaluate the quality of schools or teachers.

The pattern found among New Zealand teachers of agreement with improvement conceptions and disagreement with student accountability conceptions is consistent with, though more complex than the discredited false dichotomy between 'summative bad' and 'formative good' models of conceiving assessment—such as Carr's (2001) accountability-oriented folk model of assessment versus an improvement-oriented alternative model, Torrance and Pryor's (1998) accountability-oriented convergent assessment contrasted with teaching-improvement or divergent model of assessment, and Philipp *et al.*'s (1994) evaluation for reporting contrasted with assessment used to inform teaching. The model reported here is not framed around a simple dichotomy; rather it is multi-dimensional. Simple opposites do not explain how teachers conceive of assessment nor do they provide an adequate basis for designing teacher training or assessment policy. As Locke and Hill (2003) point out the dilemmas of teaching nowadays require that teachers exercise both accountability and formative/diagnostic conceptions of assessment.

## Implications

A number of implications related to assessment policy implementation and the design of teacher professional education and development may be drawn from this pattern of NZ teachers' conceptions of assessment. If policy and training are to be effective, they must deal with teachers' conceptions as much as they deal with declarative or procedural knowledge requirements. The implementation of any new assessment policy, tool, or practice, whether at the national or local school level, needs to take account of the complex structure of teachers' conceptions of assessment to ensure success. Kahn (2000) pointed out that teachers appeared to assimilate new assessment practices (e.g., constructivist, deep) into long-standing transmission, teacher-oriented, accountability type assessment and learning frameworks. Certainly, the implementation of new standards from professional bodies or state authorities, while well intentioned, may be reduced in effectiveness if teachers' conceptions of assessment remain unchanged or unchallenged, or if teachers remain unaware of their own conceptions. Simply introducing an assessment innovation, as in the hypothetical conversation at the start of the article, even if it is accompanied by appropriate teacher professional development, will not necessarily achieve policy objectives unless the differing, interlocked conceptions of teachers are exposed and addressed. Otherwise, quite possibly few teachers will adopt and utilize the innovation in a manner consistent with the intentions of developers of the innovation.

This research showed that teachers agreed with the conceptions that assessment

improves teaching and learning and that assessment makes schools accountable, while rejecting assessment's irrelevance in this context. Further, this positive attitude to improvement was paralleled with rejection of assessment for student accountability purposes. This suggests that the introduction of any assessment policy intended not only to diagnose and monitor student learning but also to improve the quality and quantity of learning should be done in such a way as to minimize association with student accountability and instead maximize association with teachers' commitment to improving their own instruction and the learning of their own students, while taking advantage of teachers' agreement that assessment can identify quality schooling. In other words, assessment policy may be most powerful if structured as a means of giving education professionals self-managed feedback about the quality of their own work. Emphasis on a school-based and managed process of improvement-oriented evaluation of student assessment results is likely to result in educational improvement in the quality of teaching and the quality of student learning outcomes (see for example the SEMO model, Timperley & Robinson, 2002). The implication of this research is that the development of assessment policy should include identification of and appropriate response to teachers' conceptions of assessment.

Likewise, teacher professional pre-service preparation and in-service development in the area of assessment needs to take account of teachers' pre-existing conceptions, if it is to be effective in moving teachers toward a desired set of conceptions. Pajares (1992) points out that the origins of teachers' conceptions is their own experience as students of assessment; if that experience is largely one of accountability and irrelevance then it is likely that some teacher trainees and practising teachers will have conceptions that need to be developed. As a case in point, in one small school of ten teachers including the principal, it was found that the principal agreed strongly with the improvement conception and disagreed with the irrelevance conception. In contrast, three of the teachers had much higher scores on irrelevance and accountability conceptions and disagreed with the improvement conception. Fundamentally, despite talking about the common word 'assessment' these teachers and their principal were talking past each other. An improvement-oriented assessment policy or practice in that school, without explicit attention to the differing conceptions of assessment held by the teachers, would likely be adopted and assimilated into the pre-existing conception of assessment as something that may be used but ignored.

The COA-III instrument may be useful in obtaining information to plan policy and to assist in professional development. Use in a training context may make more explicit the conceptions participants have about assessment and trigger discussions leading to appropriate selection of conceptions. Thus, the COA-III instrument could be a powerful tool in shaping assessment policy and teacher professional development in such a way that assessment improves the quality of teaching and raises the standards of student achievement.

## **Acknowledgement**

The research was conducted as part of the author's doctoral dissertation in Education

at the University of Auckland with ethical approval from the University of Auckland Human Subjects Ethics Committee (No. 2001/Q/001). The author would like to acknowledge the financial assistance of Auckland UniServices Ltd., the supervision of Professor John Hattie, and the co-operation of hundreds of teachers and trainees. The constructive comments of two anonymous reviewers are also appreciated.

### Notes on contributor

Gavin Brown is the senior project manager for the Assessment Tools for Teaching and Learning Project ([www.asttle.org.nz](http://www.asttle.org.nz)) at the University of Auckland. He has worked as an assessment researcher at the New Zealand Council for Educational Research and the University of Auckland. His research interests include large-scale assessment, information skills, structural equation modelling, and literacy.

### References

- Airasian, P. W. (1997) *Classroom assessment* (New York, McGraw-Hill).
- Asch, R. L. (1976) Teaching beliefs and evaluation, *Art Education*, 29(6), 18–22.
- Black, P. & Wiliam, D. (1998) Assessment and classroom learning, *Educational Assessment: Principles, Policy and Practice*, 5(1), 7–74.
- Borko, H., Mayfield, V., Marion, S., Flexer, R. & Cumbo, K. (1997) Teachers' developing ideas and practices about mathematics performance assessment: successes, stumbling blocks, and implications for professional development, *Teaching & Teacher Education*, 13(3), 259–278.
- Brown, G. T. L. (2002) *Teachers' conceptions of assessment* (Auckland, NZ, University of Auckland).
- Brown, G. T. L. (2004) Measuring attitude with a positively packed self-report agreement rating scale, *Psychological Reports*, 94, 1015–1024.
- Butterfield, S., Williams, A. & Marr, A. (1999) Talking about assessment: mentor-student dialogues about pupil assessment in initial teacher training, *Assessment in Education*, 6(2), 225–246.
- Calderhead, J. (1996) Teachers' beliefs and knowledge, in: D. C. Berliner & R. C. Calfee (Eds) *Handbook of educational psychology* (New York, Simon & Schuster Macmillan), 709–725.
- Carr, M. (2001) *Assessment in early childhood settings: learning stories* (London, Paul Chapman).
- Cizek, G. J., Fitzgerald, S., Shawn, M. & Rachor, R. E. (1995) Teachers' assessment practices: preparation, isolation and the kitchen sink, *Educational Assessment*, 3, 159–179.
- Clark, C. & Peterson, P. (1986) Teachers' thought processes, in: M. Wittrock (Ed.) *Handbook of research on teaching*. (3rd edn) (New York, MacMillan), 255–296.
- Cohen, D. K. & Hill, H. C. (2000) Instructional policy and classroom performance: The mathematics reform in California, *Teachers College Record*, 102, 294–343.
- Cooper, P. & Davies, C. (1993) The impact of national curriculum assessment arrangements on English teachers' thinking and classroom practice in English secondary schools, *Teaching & Teacher Education*, 9, 559–570.
- Croft, A. C. & Reid, N. A. (1991) *How often and for what purposes are NZCER tests used in primary and secondary schools?* (Wellington, NZ, NZCER).
- Croft, A. C., Strafford, E. & Mapa, L. (2000) *Stocktake/evaluation of existing diagnostic tools in literacy and numeracy in English* (Wellington, NZ, NZCER).
- Crooks, T. (1988) The impact of classroom evaluation practices on students, *Review of Educational Research*, 58(4), 438–481.
- Crooks, T. J. (2002) Educational assessment in New Zealand schools, *Assessment in Education: Principles Policy & Practice*, 9(2), 237–253.

- Dixon, H. (1999) *The effect of policy on practice: an analysis of teachers' perceptions of school based assessment practice* (Albany, NZ, Massey University).
- Dunn, K., Strafford, E. & Marston, C. (2003) *Classroom assessment practices in English and mathematics at Years 5, 7, and 9* (Wellington, NZ, NZCER).
- Ekeblad, E. & Bond, C. (1994) The nature of a conception: questions of context, in: R. Ballantyne & C. Bruce (Eds) *Phenomenography: philosophy and practice* (Brisbane, Queensland University of Technology, Centre for Applied Environmental and Social Education Research), 343–353.
- Firestone, W. A., Mayrowetz, D. & Fairman, J. (1998) Performance-based assessment and instructional change: the effects of testing in Maine and Maryland, *Educational Evaluation and Policy Analysis*, 20(2), 95–113.
- Fiske, E. B. & Ladd, H. F. (2000) *When schools compete: a cautionary tale* (Washington DC, Brookings Institution Press).
- Gipps, C., Brown, M., McCallum, B. & McAlister, S. (1995) *Intuition or evidence? Teachers and national assessment of seven-year-olds* (Buckingham, Open University Press).
- Guthrie, J. T. (2002) Preparing students for high-stakes test taking in reading, in: A. E. Farstrup & S. J. Samuels (Eds) *What research has to say about reading instruction* (Newark, DE, International Reading Association), 370–391.
- Harootunian, B. & Yarger, G. P. (1981) *Teachers' conceptions of their own success: current issues* (Washington DC, ERIC Clearinghouse on Teacher Education).
- Hawley, W. D. & Valli, L. (1999) The essentials of effective professional development: A new consensus, in: L. Darling-Hammond & G. Sykes (Eds) *Teaching as the learning profession: handbook of policy and practice* (San Francisco, CA, Jossey-Bass), 127–150.
- Heaton, J. B. (1975) *Writing English language tests* (London, Longman).
- Hershberg, T. (2002) Comment, in: D. Ravitch (Ed.) *Brookings Papers on Education Policy: 2002*. (Washington DC, Brookings Institution Press), 324–333.
- Hill, M. F. (2000) *Remapping the assessment landscape: primary teachers reconstructing assessment in self-managing schools* (Hamilton, NZ, University of Waikato).
- Hoyle, R. H. (1995) The structural equation modeling approach: basic concepts and fundamental issues, in: R. H. Hoyle (Ed.) *Structural equation modeling: concepts, issues, and applications* (Thousand Oaks, CA, Sage), 1–15.
- Kahn, E. A. (2000) A case study of assessment in a grade 10 English course, *The Journal of Educational Research*, 93, 276–286.
- Levin, B. (2001) *Reforming education: from origins to outcomes* (London, RoutledgeFalmer).
- Linn, R. L. (2000) Assessments and accountability, *Educational Researcher*, 29(2), 4–16.
- Linn, R. L. & Gronlund, N. E. (2000) *Measurement and evaluation in teaching* (New York, Macmillan).
- Locke, T. & Hill, M. F. (2003) *The impact of changes in the nature of teachers' work on teacher professionalism* (Hamilton, NZ, The University of Waikato).
- Marton, F. (1981) Phenomenography—describing conceptions of the world around us, *Instructional Science*, 10, 177–200.
- Maruyama, G. M. (1998) *Basics of structural equation modeling* (Thousand Oaks, CA, Sage).
- McMillan, J. H. (2001) *Classroom assessment: principles and practice for effective instruction* (Boston, MA, Allyn & Bacon).
- McMillan, J. H., Myran, S. & Workman, D. (2002) Elementary teachers' classroom assessment and grading practices, *The Journal of Educational Research*, 95(4), 203–213.
- Mehrens, W. A. & Lehmann, I. J. (1984) *Measurement and evaluation in education and psychology* (New York, Holt, Rinehart and Winston).
- Ministry of Education (1993) *The New Zealand curriculum framework: te anga marautanga o Aotearoa* (Wellington, NZ, Learning Media).
- Ministry of Education (1994) *Assessment: policy to practice* (Wellington, NZ, Learning Media).

- New Zealand (1998) *Assessment for success in primary schools: green paper* (Wellington, NZ, Ministry of Education).
- Noble, A. J. & Smith, M. L. (1994) *Old and new beliefs about measurement-driven reform: 'The more things change, the more they stay the same'* (Los Angeles, CA, UCLA, CRESST).
- Pajares, M. F. (1992) Teachers' beliefs and educational research: cleaning up a messy construct, *Review of Educational Research*, 62, 307–332.
- Philipp, R.A., Flores, A., Sowder, J. T. & Schappelle, B. P. (1994) Conceptions and practices of extraordinary mathematics teachers, *Journal of Mathematical Behavior*, 13, 155–180.
- Popham, W. J. (2000) *Modern educational measurement: practical guidelines for educational leaders* (Boston, MA, Allyn & Bacon).
- Pratt, D. D. (1992) Conceptions of teaching, *Adult Education Quarterly*, 42(4), 203–220.
- Quilter, S. M. (1998) *Inservice teachers' assessment literacy and attitudes toward assessment* (Columbia, SC, University of South Carolina).
- Resnick, L. B. & Resnick, D. P. (1989) *Assessing the thinking curriculum: new tools for educational reform* (Washington DC, National Commission on Testing and Public Policy).
- Smith, M. L. & Fey, P. (2000) Validity and accountability in high-stakes testing, *Journal of Teacher Education*, 51(5), 334–344.
- Smith, M. L., Heinecke, W. & Noble, A. J. (1999) Assessment policy and political spectacle, *Teachers College Record*, 101(2), 157–191.
- Sturrock, F. (1999) *Teacher census: preliminary report* (Wellington, NZ, Ministry of Education, Demographic and Statistical Analysis Unit).
- Thompson, A. G. (1992) Teachers' beliefs and conceptions: A synthesis of the research, in: D. A. Grouws (Ed.) *Handbook of research on mathematics teaching and learning* (New York, Macmillan), 127–146.
- Timperley, H. & Robinson, V. (2002) *Partnership: focusing the relationship on the task of school improvement* (Wellington, NZ, NZCER).
- Tittle, C. K. (1994) Toward an educational psychology of assessment for teaching and learning: theories, contexts, and validation arguments, *Educational Psychologist*, 29, 149–162.
- Torrance, H. & Pryor, J. (1998) *Investigating formative assessment: teaching, learning and assessment in the classroom* (Buckingham, Open University Press).
- van den Berg, B. (2002) Teachers' meanings regarding educational practice, *Review of Educational Research*, (72), 577–625.
- Warren, E. & Nisbet, S. (1999) The relationship between the purported use of assessment techniques and beliefs about the uses of assessment, in: J. M. Truran & K. M. Truran (Eds) *22nd Annual Conference of the Mathematics Education and Research Group of Australasia* (Adelaide, SA, MERGA), 515–521.
- Webb, N. L. (1992) Assessment of students' knowledge of mathematics: steps toward a theory, in: D. A. Grouws (Ed.) *Handbook of research on mathematics teaching and learning* (New York, Macmillan), 661–683.
- Wylie, C. (1997) *Self-managing schools seven years on: what have we learnt?* (Wellington, NZ, NZCER).